DOCUMENT RESUME

ED 165 626 HE 010 889

AUTHOR Costello, Francis J.
TITLE An Analysis of a Student Evaluation of Instruction in the Graduate School of an Urban University.
PUB DATE 77
NOTE 36p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Class Size; *College Faculty; College Students; Evaluation Methods; *Grades (Scholastic); Graduate Students; Higher Education; Questionnaires; Rating Scales; *Reliability; *Student Evaluation of Teacher Performance; Student Opinion; *Student Teacher Relationship; *Teacher Evaluation; Urban Universities
IDENTIFIERS *University of New Haven CT

ABSTRACT

Since the administrative evaluation of faculty at the University of New Haven has been hampered by the faculty's distrust of student input, this study aimed to improve the faculty's confidence in student evaluations of instruction. A student evaluation of instructors was statistically compared to grades issued by the instructors, and to class size. The evaluation was prepared by graduate students at the University of New Haven. Results of the analysis showed no statistical relationship between grades issued by an instructor and the rating the instructor received from the students. There was also no statistical relationship between the instructor's rating and class size. Grades issued by evaluated instructors were statistically compared to grades issued by instructors who were not evaluated, and no statistical difference was found. It is recommended that university-wide student evaluations be used for a two-year trial period for use in the evaluation of faculty by the university administration. Negative use of the evaluation results by the administration is not recommended. Statistical data and the survey instrument are appended. (Author/LBH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED165626

AN ANALYSIS OF A STUDENT EVALUATION OF INSTRUCTION IN THE GRADUATE SCHOOL OF AN URBAN UNIVERSITY

BY

FRANCIS J. COSTELLO Ed. D
NOVA UNIVERSITY
UNIVERSITY OF NEW HAVEN
1977

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Francis J. Costello

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

HE 010 559

ABSTRACT
1977

An Analysis of a Student Evaluation of Instruction

F. J. Costello

The administrative evaluation of faculty at the University of New Haven has been hampered by the faculty's distrust of student input. The purpose of this study was to improve the faculty's confidence in student evaluations of instruction.

A student evaluation of instructors was statistically compared to grades issued by the instructors, and to class size. The evaluation was prepared by students in the graduate school at the University of New Haven.

Results of the analysis showed no statistical relationship between grades issed by an instructor and the rating the instructor received from the students. Also, there was no statistical relationship between the instructor's rating and the class size.

Grades issued by evaluated instructors were statistically compared to grades issued by instructors who were not evaluated. There was no statistical difference between the grades.

Recommendations suggest that University wide student evaluations be used for a two year trial period for use in the evaluation of faculty by the University administration. Negative use of the evaluation results by the administration was not recommended.

TABLE OF CONTENTS

Context of problem	1
Statement of the problem	3
Review of the literature	5
Significance	12
Statement of hypothesis	13
Methods and proceedures	14
Results	16
Discussion	17
Conclusions and recommendations	19
References	22
Appendix 1	25
Appendix 2	27
Appendix 3	29
Appendix 4	31

An Analysis of a Student Evaluation of Instruction

CONTEXT OF THE PROBLEM

"It is no exaggeration to say that the most critical problem confronting any university is the proper evaluation of faculty." The quote is by L. Wilson, president of the University of Texas in 1942.(1)

"Today, faculty evaluation is one of the most complex aspects of the academic world." This quote is by R. I. Miller, made in 1972. (2)

Over the years, the literature covering ways and means of evaluating college faculty has developed extensively; yet, the problem of implementing and administering a successful evaluation program at a specific institution is almost insurmountable.

Many factors that are commonly used to evaluate college teachers are factual in nature. Such information as the number of degrees, number of years teaching, published papers, etc., present little or no problem since they are relatively non-critical. However, the question of whether or not the teacher is a "good" teacher presents a problem since the answer involves judgment and opinion. Moreover, who is to answer the question? Logically , the student who receives the product should be the judge of the effectiveness of the product. Consequently, evaluation programs at colleges and universities usually include student input.

Professors, like anyone else, stiffen at the prospect of being evaluated, especially if the evaluation is by students.

An adverse student rating can destroy a teacher's confidence and promote defensiveness that would be destructive to effective teaching. Moreover, the results of a student evaluation inevitably find there way into the administrative offices where tenure, merit increases, and promotion are subject to approval or opposition. (3) (4)

The main fear that the professor has relative to a student evaluation is that the appraisal is directly related to the grades issued by the teacher. Good grades produce good evaluations.

The proposition that student appraisals of teachers can be a means of retaliation by students who are doing C work is supported by some research. (5) However, other research has found student ratings to be more objective and perhaps more valid than alternative evaluation procedures. (6)

It appears that the usefullness of student evaluations hinge on whether or not the appraisal response is valid and reliable. However, it is questionable as to whether or not this type of evaluation can achieve reliability and validity. (7)

STATEMENT OF THE PROBLEM

Faculty evaluation has been an objective at the University of New Haven for several years. In 1976 the administration issued to all departments a directive requesting evaluations of all faculty personnel. The results of the forced evaluations was, at best, a resume prepared by the individual faculty person. Many departments had meetings and decided to ignore the directive. The main reason given for the failure of the evaluation effort was that the teacher's effectiveness in the classroom could not be objectively and reliably determined.

Student evaluations have been used at the University a few times in the past. The evaluations were sponsored by the Student Union. However, many professors would not allow the evaluation forms to be used in his class, and the Student Union was reluctant to force an evaluation without the permission of the instructor. The professors at the University generally agree that the student's evaluation of an instructor is merely a reflection of the grades issued by that instructor. Extreme cases are usually cited to prove the grade-rating relationship. A few instructors who are noted for issuing nothing but A-B grades invariably receive a very high student rating while a few teachers, noted as "task masters" receive low ratings.

If some evidence of credibility could be attached to the student's evaluation efforts, perhaps the faculty would be more inclined to accept the student's rating system.

Moreover, if the faculty developed some degree of confidence in the student evaluation system, the administration's objective of faculty evaluations would have a better chance of being accomplished.

The professional view must consider the proposition that the need for faculty evaluation includes improved learning as well as promotion, tenure and merit increases in salary. (8) Also, impartial student input is essential to the complete evaluation of faculty since student input seems to be the logical way of determining the attitude and performance of a teacher in a particular classroom.

The purpose of this study is to determine whether or not a recent evaluation of teachers by students at the University of New Haven are related to the grades issued by those teachers. Moreover, the purpose is to give the faculty statistical evidence of the relationship between ratings and grades in an effort to relax the faculty's anxiety relative to student appraisals and to effectively promote the implementation of the evaluation of the faculty with student input.

HE 010 889

ABST::

Since the administrative evaluation of faculty at the University of New Haven has been hampered by the faculty's distrust of student input, this study aimed to improve the faculty's confidence in student evaluations of instruction. A student evaluation of instructors was statistically compared to grades issued by the instructors, and to class size. The evaluation was prepared by graduate students at the University of New Haven. Results of the analysis showed no statistical relationship between grades issued by and instructor and the rating the instructor received from the students. There was also no statistical relationship between the instructor's rating and class size. Grades issued by evaluated instructors were statisically compared to grades issued by instructors who were not evaluated, and no statistical difference was found. It is recommended that university-wide student evaluations be used for a two-year trial period for use in the evaluation of faculty by the university administration. Negative use of the evaluation results by the administration is not recommended. Statistical data and the survey instrument are appended. (Author/LBH)

IDEN:: *University of New Haven CT

DESC::

Higher Education; *Student Evaluation of Teacher Performance; *Teacher Evaluation; *College Faculty; *Grades (Scholastic); Evaluation Methods; Class Size; Questionnaires; Rating Scales; *Reliability; Student Opinion; *Student Teacher Relationship; College STudents; Graduate Students; Urban Universities

REVIEW OF THE LITERATURE

A number of methods, either singularly or in combination, seem to be in use today by educational institutions for the purpose of evaluating individual faculty members. Regardless of the method, the purpose of evaluation should be to improve learning, and the critical measurement that is required for effective and ovjective evaluation is teaching effectiveness.

However, the determination of a teacher's effectiveness in the classroom is largely determined by the judgment of chairman, deans, and colleagues. The source of information that is most likely to provide data on the faculty member's effectiveness as a teacher is seldom used. That source is the students. (9)

Classroom Visitation

The classroom visit, for purposes of appraisal, by a dean, chairman or colleague is an obvious way of obtaining information on the teacher's effectiveness in teaching. The system however, is seldom used. Objections by faculty include freedom of the classroom and infringement of the professional dignity of the teacher. (10)

There have been some notable experiments in the classroom visitation appraisals of teachers. At Carnegie-Mellan University, classroom appraisals of teachers was introduced in 1969 and 1970. Visits to classrooms by senior faculty members were carried out under

strong protest by the faculty. The plan was ill-conceived since the visits were conducted by selected senior faculty and reciprocal visits were not part of the plan. Fear of deviciveness and incompetentcy of the appraisor plagued the junior faculty. Nevertheless, some useful results were realized by the Carnegie program. (11)

Self Appraisal

Self evaluation is constantly in the mind of professionals, although this tendency to self evaluate seldom becomes systematic or orderly, and seldom gets onto paper. The self evaluation for appraisal purposes presents a comprehensive check sheet to the teacher who rates himself and thereby merely organizes his previous evaluative thinking.

One of the problems with self evaluation is the tendency on the part of certain types of teachers to be over-critical or over-protective in making the appraisal.

The usual faculty objection to evaluation is lessened when self evaluation is used, and the tendency to associate the evaluation with improved teaching is enhanced. (12) (13)

Course Material and Data

Another method of evaluating teaching effectiveness is the materials and procedure approach. This system is often used in

engineering schools and the writer has had nine years of experience with the "course folder" procedure.

Each course has a syllabus that specifies the objective of the course in terms of the specific material to be taught, the text to be used and a bibliography. At the end of the semester the instructor presents to the department an outline of his course. The outline specifies the material covered in class, examples of student's test papers, and the final grade spread. Examples of student's term projects may also be included.

The department chairman will review the course folders and compare the requirements for the course with the instructor's outline of what was taught. Samples of student's tests will confirm, at least to a fair degree, that the material was covered by the instructor. Also, the nature of the test problems and the quality of the student's responses will confirm a degree of quality relative to the instruction and the amount of learning that took place.

Accreditation of engineering schools by the Engineer's Council on Professional Development is based in part on an examination of syllabi and course folders. The accrediting review team will pass judgment on the quality of instruction by direct examination of the course folders that must be maintained for a seven year period.

Publications and Other Activity

The evaluation of a teacher's ability or competence may be narrowed down to the number of publications he has managed to produce over a period of time. Whether or not the publication is "prestigeous" depends on the vehicle of publication. A scientific journal is most prestigeous while a trade magazine carries little weight.

A tremendous amount of garbage is published every year. Nevertheless, garbage publications add "yardage" to the teacher's resume. The resume may be read or counted and in either case, the desired effect may be achieved. (14)

Included in the resume type evaluation are activities such as committee work, activity in professional associations and community involvement.

Visual and Audio Devices

Some work has been done at Harvard by K. Morimoto in evaluating teaching effectiveness by using recording devices. The project has not been reported thoroughly, but there is some indication that it is an effective way of increasing the effectiveness of the teacher.

It is interesting to note that industrial engineers have for many years used motion picture analysis to increase the effectiveness of workers on assembly lines.

Student Opinion

Perhaps the most controversial method of evaluating a teacher is the use of student opinion. Common points of criticism and support of student evaluations are summarized by K. E. Eble (3) and are presented here.

Criticism:

1) Student evaluations raise the anxiety of the faculty and thereby interfere with the learning process.

2) Student evaluations amount to a popularity contest and force the faculty into competition.

3) The systems are rigid and introduce further mechanization of education.

4) The information furnished by the students are inadequate and misleading.

5) The effect of the student's evaluation is short term in nature and does not effect the basic functioning of the institution as a whole.

6) Cooperation between faculty and students is strained.

Support:

1) Student evaluations increase the chance that excellent teaching will be rewarded.

2) Student participation in the entire learning process will be

enhanced.

3) The system provides the only first hand information about a teacher's classroom performance.

4) It may help to stimulate the institution to place teaching competence as a main goal.

5) Faculty and administration support of student evaluation is an outward sign of confidence in student involvement in the institution's goals and practices.

There is rather little to be found in the literature relative to student evaluation of teachers that involved data taking on questionaires that were backed up with information on reliability. One paper was found in the literature that presented a questionaire that was supported by evidence of validity. (19)

Some work on the construction of a reliable questionaire was done by S. L. Whiteman of the Baldwin-Wallace College in Kentucky. The work was done in 1970-1971. He found, using his questionaire, a test-retest reliability of .68 (product moment correlation). However, the proposition of validity is in doubt. His research also revealed that:

1) Instructors of small classes received higher ratings than instructors of large classes.

2) Students who were told that the instructor was "good" rated him high while students who were told that the instructor was

"bad" rated him low.

3) Upper classmen generally rated instructors high while lower classmen generally rated him lower.

4) Students who expressed a genuine interest in the subject matter gave high ratings while students who expressed "no particular interest" rated the instructor low.

Relative to the findings listed above, it seems that a student's evaluation of a teacher is effected by quite a few factors that are not directly or fairly related to the teacher's competence in the classroom.

Further doubt of the reliability and validity of student's appraisal of faculty can be entertained when the question of student competence is taken into consideration. It is questionable if the student body has the knowledge, experience and ability to evaluate a professor with accuracy and fairness. (21) Moreover, student ratings may only identify negligence in a professor. The rating may not be a valid tool to be used for judgments on teaching ability in a university since a university is a place of learning, not a place of teaching. The university students are there to learn, not to be taught. (22)

Regardless of the questions on validity and reliability, there seems to be sufficient research findings to support the proposition that student ratings of instructors are not related to the grades issued by

the particular instructor.

Blum's study at the City College of New York in 1936 came to the conclusion that student's actual or estimated grades did not influence the evaluation of the teacher. (23) In 1953, Amhuff, via a study that involved 1500 samples, determined that there was no relationship between the grades received by upper classmen and the rating of faculty by those students. (24) Clark and Keller, in a 1954 study at the University of Minnesota, found that bright students produced ratings of faculty that were the same as ratings produced by poorer students. (25)

Again, in 1972, J. A. Colliver, as a result of a study at the University of Illinois, reported that there was no relationship between grades and student's evaluation of faculty. (28)

Finally, in 1974, D. S. Holmes conducted a study at the University of Texas. Over 700 evaluations were compared to the student's self expressed expected grade. There was no relationship between the student's rating of his instructor and the grade the student expected to receive. (29)

SIGNIFICANCE

At the University of New Haven, student evaluation of teachers has been strongly resisted by many of the faculty. The main concern

of the faculty is that the student's evaluation is directly related to the grades that an instructor is in the habit of issuing. However, no factual analysis of this question has ever been performed, consequently the issue remains in the debate stage with spotty evidence or "examples" being presented from time to time by both sides in the controversy.

This issue may be put to rest, one way or the other, by a relatively simple statistical analysis.

If credibility can be attached to the student's evaluation of faculty at the University, the information may become part of the administration's evaluative procedures and, consequently, greater emphasis will be attached to competence in teaching and, learning may be enhanced.

STATEMENT OF HYPOTHESIS

1) The student's rating of instructors is not related to the grades the students received in the course.

2) Grades issued by instructors who did not permit evaluation, are the same as grades issued by instructors who allowed the students to run the evaluation.

3) Class size is inversely related to the instructor's rating.

RATIONALE FOR THE HYPOTHESIS

Some research has shown a relationship between the rating a student gives his teacher and the grade the student receives. However, the bulk of the available literature indicates that there is no relationship between the rating and the grade. (15, 17, 18) Moreover, the non-relationship is supported by research that identifies factors other than grades as having a major influence on the rating. (16, 20)

There is no reason to suspect that students at the University of New Haven should react to the rating in a manner that is significantly different from students in other schools.

METHODS AND PROCEDURES

An evaluation of faculty was performed by the students in the graduate school during the fall semester of 1976. The evaluation was administered on the last day of class. Evaluation forms (see appendix 4) were distributed by the instructor and the completed forms were collected and sealed in an envelope by the students in the class.

One hundred fifteen classes were conducted during the semester. Twenty-one instructors were not evaluated because the instructor did not distribute the evaluation forms. Evaluations were received on the remaining 94 instructors. Each instructor was evaluated via 12

questions, each of which was graded by the student from 1 to 5, with 5 being the highest rating. The results of the evaluation were published.

Grades issued by instructors were available through the registrar.

An average rating for each instructor was calculated by adding the ratings of the 12 questions and dividing by 12. The average grade issued by an instructor was calculated by using the grade point average system with an A carrying 4 points.

Hypothesis 1 was tested via a random sample of 31 instructors drawn from the 94 who were evaluated. An average rating and the average grade issued by each instructor in the sample was calculated. A Pearson correlation factor was calculated and statistically tested to 95% confidence.

Hypothesis 2 was tested by comparing the average grades issued by the 31 instructors used in checking hypothesis 1, to the average grades issued by the 21 instructors who were not evaluated. The difference between the mean grades of the two groups was tested for significance via a t statistic, to 95% confidence.

Hypothesis 3 deals with the proposition that class size is inversely related to the instructor's rating. The 31 classes selected relative to hypothesis 1 were used to test hypothesis 3. Class sizes

were taken from the registrar's files. A Pearson correlation factor was calculated using class size and the instructor's average rating. Significance was tested to 95% confidence.

RESULTS

Hypothesis 1:

The student's rating of instructors is not related to the grades the students received in the course.

Relative to the sample of 31 classes, the mean rating was 3.94 (on a scale of 1 to 5) with a standard deviation of .489. The mean grade was 3.32 (on a scale of 0 to 4) and the standard deviation was .3009. The Pearson correlation factor (r) is .064. The "r" factor of .064 will not support significance to 95% confidence, therefore, the hypothesis is accepted. (For data and calculations see Appendix 1)

Hypothesis 2:

Grades issued by instructors who did not permit an evaluation are the same as grades issued by instructors who were evaluated.

The mean grade issued by instructors who were evaluated (sample size 31) was 3.32 with a standard deviation of .3009. Instructors who did not permit an evaluation of their course, 21 in number, issued a mean grade of 3.41 with standard deviation of .349. A t statistic was calculated and found to be .9917. At 95% confidence and 50 degrees

of freedom, a t of 2.01 is needed for significance, therefore the hypothesis is accepted. (For data and calculations see Appendix 2)

Hypothesis 3:

Class size is inversely related to the instructor's rating.

The average class size, using the sample of 31 evaluated courses, was 20.58 with standard deviation of 7.68. A comparison of the class sizes and the instructors' ratings produced a Pearson correlation factor of -.235. At 95% confidence, the Z required for significance is 1.64 since this is a one tailed (directional) hypothesis. At 30 degrees of freedom, the data produced a Z of 1.28, not enough to show a relationship between class size and rating. The Z of 1.28 would support significance at 90% confidence. However, the hypothesis must be rejected. (For data and calculations see Appendix 3)

DISCUSSION

For practical purposes, there is no relationship between the grades issed by an instructor, and the rating the instructor received from his students. The correlation factor of .064 that relates grades to rating, accounts for only 1/2 of one percent of the variance. This result conforms to the findings of several other studies, and agrees with Knowles conclusions.

The lack of relationship between grades issued by evaluated instructors and non-evaluated instructors is supported by a t statistic of .99 which is well below the significant level. However, instructors who did not permit an evaluation may have been under pressure to issue grades that would give no cause for alarm relative to an analysis such as this study. Yet, the t of .99 is far enough below the required 2.01 to soften suspicion. Further investigation would involve an analysis of the various instructor's grades that were issued over the past semesters, and such a specific investigation seems unwise at this time. However, a spot check of non-evaluated instructors whose grading habits are known did not reveal departures from past practice.

An indication of a relationship developed between the instructor's rating and class size. The correlation of -.233 clearly shows the trend of an inverse relationship. An "r" of -.235 would support a relationship at the 90% confidence level (single tailed). However, the low "r" factor will explain only 5% of the variance.

Other investigators have found significant relationships between class size and ratings, and the results of this study, using a larger sample, may very well have driven the "r" factor to .3, the 95% confidence level. Even at .3, only 10% of the variance in the ratings could be attributed to class size.

CONCLUSIONS

1) Student evaluations of instructors are not effected by the grades the students receive from the instructor.

2) Grades issued by instructors who do not allow student evaluations are statistically the same as grades issued by instructors who cooperate with evaluations.

3) Class size has a small inverse effect on teacher rating. Students give higher ratings when the class size is small. This conclusion is an extrapolation from the results of this study. Statistically, the conclusion that class size and ratings are related cannot be supported.

RECOMMENDATIONS

The administration should petition the faculty, via a general faculty vote, to set aside a four semester trial period that would require a student evaluation of every class that functions in the University during that period. If the faculty approved, and a majority of faculty cooperate with evaluations at the present time, the evaluation requirement could be made a part of the annual faculty-administration agreement. The annual agreement becomes part of the yearly contract for each of the faculty. Compliance would be required by the faculty.

It should be agreed that negative action, as an outcome of the

evaluations, against any instructor would be withheld. Instead, it should be recognized by all parties, students, faculty and administration, that the purpose of the evaluations is to enhance learning. Instructors who consistently receive bad ratings may be urged to take advantage of courses in higher education that are pertinent to their problem area. Several local state colleges offer courses in education.

However, a significant improvement in learning may take place as a result of instructors being made aware of apparently minor faults; the otherwise excellent instructor who at times becomes inaudible to the students, or the instructor who is highly respected by students who complain that he does not clearly explain his homework assignments. The writer is aware of one case involving a professor of economics. The professor received the highest rating on all questions except one that pertained to class participation, and the students' reaction to this question was "little or none". The professor just didn't realize that he talked continuously and the students didn't have a chance to get a word in edgewise. The instructor now, periodically during a lecture, pauses and asks for questions. Learning has been enhanced as a result of feedback via a student evaluation of instruction.

If the trial period evaluations prove to be a positive experience

to all concerned, the evaluations may become a permanent part of the practice of education at the University.

The trial evaluations should be subjected to a statistical analysis with results disseminated among the students and faculty. Efforts should be made to establish the validity and reliability of the evaluation form now in use.

LIMITATIONS

The results of this study are limited to the specific classes, students and instructors, at the University of New Haven, that participated in the 1976 graduate school evaluations. If an instructor teaches a different course, or even the same course with a different student mix, his evaluation could be entirely different from the 1976 evaluation. A single evaluation is a poor basis for making final judgments.

REFERENCES

1) L. Wilson, The Academic Man
Oxford University Press, 1942

2) R. I. Miller, Evaluating Faculty Performance
Jossey-Boss, 1972

3) K. E. Eble, The Recognition and Evaluation of Teaching
American Association of University Professors, 1972

4) F. N. Kerlinger, Student Evaluation of University Professors
School and Society, 10/1971

5) D. S. Holmes, The Effects of Disconfirmed Grades on Student's Evaluations of Instructors

6) G. L. Fahey, Student Rating of Teaching
Institute for Higher Education - Conference, University of Pittsburgh, 1970

7) R. Rowland, Can Teaching Be Measured Objectively
Improving College and University Teaching 18-1970

8) W. J. McKeochie, Research on Student Ratings of Teachers
Institute for Higher Education - Conference, University of Pittsburgh, 1970

9) A. W. Astin and C. T. Lee, Current Practice in the Evaluation and Training of College Teachers
Improving College Teaching, American Council on Education - 1967

10) C. E. Gray, The Teaching Model and Evaluation of Teaching Performance
Journal of Higher Education, 40, 1969

11) A. M. Eastman, How Visitation Came to Carnegie-Mellon University
The Bulletin of the Association of Departments - Carnegie-Mellon University - May 1969

12) R. H. Simpson, Use of Teacher Self-Evaluation for Improving Instruction
Report to the American Association of Colleges for Teacher Education, 1962

13) R. H. Simpson, Teacher Self-Evaluation
MacMillan - 1966

14) J. W. Gerstad, Policies and Practices in Faculty Evaluations
Educational Record, 7-1961

15) P. L. Dressel, Evaluation of Instruction
Journal of Farm Economics, 49, 1967

16) N. L. Gage, The Appraisal of College Teaching
Journal of Higher Education

17) M. Hildebrand, Effective University Teaching and Its Evaluation
Center for R. & D. in Higher Education, 1970

18) T. D. Langen, Student Assessment of Teaching Effectiveness
Improving College and University Teaching 14, 1966

19) W. J. McKeachie, Student Ratings of Teacher Effectiveness
American Research Journal, 8 - 1971

20) D. Newell, Evaluation of Teachers
Conference on Evaluation of Teaching and Teachers - University of Kentucky - 1967

21) Brickman, W. W., Student Evaluation of Professors
School and Society 1966, #94

22) Borgatta, E. E. Student Rating of Faculty
AAUP Bulletin, 1970 #56

23) Blum, M. L. An Investigation of the Relation Existing Between Student Grades and Their Rating of the Instructor's Ability to Teach
Journal of Educational Psychology - 1936

24) Anikeeff, A. M. Factors Effecting Evaluation of College Faculty
Journal of Applied Psychology - 1953

25) Clark and Keller, Student Ratings of College Teaching
University of Minnesota Press, 1954

26) Eckert, R. E. and Keller, R. J. Student Ratings of College Teaching
University of Minnesota Press -1954

27) A. S. Knowles, Handbook of College and University Administration
McGraw Hill 1970

28) Colliver, J. A. Student Evaluation of Faculty at Sangamon State University
University of Illinois Press 1972

29) Holmes, D. S. The Relationship Between Expected Grades and Student's Evaluation of Their Instructor
Educational and Psychological Measurement, 1974

APPENDIX 1 *

CALCULATIONS OF CORRELATION BETWEEN COURSE RATINGS AND GRADES

Course #	Mean Rating = x	Mean Grade = y
1	4.54	3.26
2	3.93	3.08
3	3.04	3.26
4	4.04	3.32
5	4.60	3.40
6	3.59	2.79
7	4.77	3.58
8	4.23	3.68
9	4.27	2.93
10	4.00	3.50
11	4.27	3.66
12	4.26	2.92
13	3.88	2.91
14	4.07	3.16
15	3.62	3.44
16	4.14	3.44
17	3.63	3.32
18	3.03	3.22
19	3.73	3.10
20	3.76	3.25
21	4.09	3.10
22	3.33	3.26
23	4.63	3.62
24	3.86	3.20
25	4.42	3.89
26	3.66	3.66
27	3.97	3.71
28	4.36	3.59
29	4.27	2.75
30	3.65	3.16
31	2.66	3.82

*NOTE: Statistical equations used in Appendix 1 through 3 were taken from:

Bruning and Kintz, Computational Handbook of Statistics
Scott-Foresman, 1968

$$r = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{\left\{\left[N\Sigma x^2 - (\Sigma x)^2\right]\left[N\Sigma y^2 - (\Sigma y)^2\right]\right\}^{\frac{1}{2}}}$$

$$r = \frac{31(406.55) - (122.3)(102.98)}{\left\{\left[31(489.59) - (122.3)^2\right]\left[31(344.72) - (102.98)^2\right]\right\}^{\frac{1}{2}}} = .064$$

APPENDIX 2

CALCULATION OF t STATISTIC
GRADES OF EVALUATED AND NON-EVALUATED INSTRUCTORS

X^1 Grades : Evaluated Instructors	X^2 Grades : Non-Evaluated Instructors
3.26	3.42
3.08	3.22
3.26	3.55
3.32	3.40
3.40	3.00
2.79	3.35
3.58	3.82
3.68	3.88
2.93	2.37
3.50	3.14
3.66	3.31
2.92	3.57
2.91	3.77
3.16	3.35
3.44	3.73
3.44	3.35
3.32	3.93
3.22	3.16
3.10	3.56
3.25	3.54
3.10	3.27
3.26	
3.62	
3.20	
3.89	
3.66	
3.71	
3.59	
2.75	
3.65	
2.66	

$N = 31$	$N = 21$
$\bar{X}_1 = 3.32$	$\bar{X}_2 = 3.41$
$S = .3009$	$S = .349$
$\Sigma X = 102.98$	$\Sigma X = 71.69$
$\Sigma X^2 = 344.81$	$\Sigma X^2 = 247.17$

$$\tau = \frac{\bar{X}_2 - \bar{X}_1}{\left\{\left[\frac{\Sigma X_1^2 - \frac{(\Sigma X_1)^2}{N_1} + \Sigma X_2^2 - \frac{(\Sigma X_2)^2}{N_2}}{N_1 + N_2 - 2}\right]\left[\frac{1}{N_1} + \frac{1}{N_2}\right]\right\}^{\frac{1}{2}}}$$

$$\tau = \frac{3.41 - 3.32}{\left\{\left[\frac{247.17 - \frac{5139.45}{21} + 344.81 - \frac{10604.88}{31}}{21 + 31 - 2}\right]\left[\frac{1}{21} + \frac{1}{31}\right]\right\}^{\frac{1}{2}}} = .9917$$

APPENDIX 3

CORRELATION OF COURSE RATING TO CLASS SIZE

Course #	Course Rating = x	Class Size = y
1	4.54	23
2	3.93	14
3	3.04	25
4	4.04	27
5	4.60	5
6	3.59	27
7	4.77	21
8	4.23	22
9	4.27	19
10	4.00	21
11	4.27	12
12	4.26	14
13	3.88	30
14	4.07	24
15	3.62	9
16	4.14	34
17	3.63	29
18	3.03	23
19	3.73	30
20	3.76	34
21	4.09	20
22	3.33	16
23	4.63	11
24	3.86	30
25	4.42	21
26	3.66	7
27	3.97	15
28	4.36	22
29	4.27	12
30	3.65	19
31	2.66	22

$$r = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\left\{\left[N\Sigma X^2 - (\Sigma X)^2\right]\left[N\Sigma Y^2 - (\Sigma Y)^2\right]\right\}^{\frac{1}{2}}}$$

$$r = \frac{31(2490.64) - 122.3(638)}{\left\{\left[31(489.59) - (122.3)^2\right]\left[31(14904) - (638)^2\right]\right\}^{\frac{1}{2}}} = -.235$$

APPENDIX 4

UNIVERSITY OF NEW HAVEN GRADUATE STUDENT COUNCIL COURSE EVALUATION

COURSE TITLE AND #_____ INSTRUCTOR_____

PLEASE ANSWER THE FOLLOWING QUESTIONS BY CIRCLING A NUMBER.

1. How knowledgeable was the teacher about this subject?

   Ignorant 1 2 3 4 5 Brilliant

2. How well did the teacher explain ideas and concepts?

   Utterly Confused Students 1 2 3 4 5 Explained With Perfect Clarity

3. How concerned was the instructor about student progress?

   Total Unconcerned 1 2 3 4 5 Very Concerned

4. Evaluate the instructor on the following characteristics:

   A. Open minded about other viewpoints and ways of thinking

      Closed Minded 1 2 3 4 5 Open Minded

   B. Usually well prepared for class

      Never 1 2 3 4 5 Always

5. How interesting were the lectures?

   Terribly Boring 1 2 3 4 5 Fascinating

6. How informative were the class sessions?

   Worthless 1 2 3 4 5 Very Informative

7. How adequate and pertinent was the assigned reading material?

   Unsuitable 1 2 3 4 5 Completely Suitable

8. How closely did the course content follow the catalog description?

   Unrelated 1 2 3 4 5 Exactly As Described

9. Would you recommend this course to other students?

   Never 1 2 3 4 5 Definitely

10. Would you recommend this instructor to other students?

    Never 1 2 3 4 5 Definitely

11. What part of your $$ investment do you feel was returned to you in this class?

    $$ Wasted 1 2 3 4 5 $$ Well Spent

12. About how many hours per week (average) did you devote to this course outside class? _____